REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
JANUARY 2009	Conference Paper Postprint	January 2008 – August 2008

4. TITLE AND SUBTITLE
WIRELESS VISUAL SENSOR NETWORK RESOURCE ALLOCATION USING CROSS-LAYER OPTIMIZATION

5a. CONTRACT NUMBER
IN HOUSE

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
62702F

6. AUTHOR(S)
Elizabeth S. Bentley, John D. Matyjas, Michael J. Medley, and Lisimachos P. Kondi

5d. PROJECT NUMBER
ANCL

5e. TASK NUMBER
62

5f. WORK UNIT NUMBER
07

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFRL/RIGE
525 Brooks Road
Rome, NY 13441-4505

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFRL/RIGE
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
N/A

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RI-RS-TP-2009-56

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution unlimited PA# 88ABW-2008-0320 Date Cleared: 29-September-2008

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We propose an approach to manage network resources for a Direct Sequence Code Division Multiple Access (DS-CSMA) visual sensor network where nodes monitor scenes with varying levels of motion. It uses cross-layer optimization across the physical layer, the link layer and the application layer. Our technique simultaneously assigns a source coding rate, a channel coding rate, and a power level to all nodes in the network based on one of two criteria that maximize the quality of video of the entire network as a whole, subject to a constraint on the total chip rate. One criterion results in the minimal average end-to-end distortion amongst all nodes, while the other criterion minimizes the maximum distortion of the network. Our approach allows one to determine the capacity of the visual sensor network based on the number of nodes and the quality of video that must be transmitted. For bandwidth-limited applications, one can also determine the minimum bandwidth needed to accommodate a number of nodes with a specific target chip rate.

15. SUBJECT TERMS
Cross-layer, visual sensor network, Code Division Multiple Access (CDMA), resource allocation, H.265, spread spectrum, joint source-channel coding, convolutional codes, power control, multimedia communications.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Michael J. Medley
U	U	U	UU	12	19b. TELEPHONE NUMBER (*Include area code*) N/A

Wireless Visual Sensor Network Resource Allocation using Cross-Layer Optimization

Elizabeth S. Bentley, John D. Matyjas, Michael J. Medley
Air Force Research Laboratory/RIGE, 525 Brooks Road, Rome, NY 13441
and
Lisimachos P. Kondi
University of Ioannina, Department of Computer Science, 45110 Ioannina, Greece

ABSTRACT

In this paper, we propose an approach to manage network resources for a Direct Sequence Code Division Multiple Access (DS-CDMA) visual sensor network where nodes monitor scenes with varying levels of motion. It uses cross-layer optimization across the physical layer, the link layer and the application layer. Our technique simultaneously assigns a source coding rate, a channel coding rate, and a power level to all nodes in the network based on one of two criteria that maximize the quality of video of the entire network as a whole, subject to a constraint on the total chip rate. One criterion results in the minimal average end-to-end distortion amongst all nodes, while the other criterion minimizes the maximum distortion of the network. Our approach allows one to determine the capacity of the visual sensor network based on the number of nodes and the quality of video that must be transmitted. For bandwidth-limited applications, one can also determine the minimum bandwidth needed to accommodate a number of nodes with a specific target chip rate. Video captured by a sensor node camera is encoded and decoded using the H.264 video codec by a centralized control unit at the network layer. To reduce the computational complexity of the solution, Universal Rate-Distortion Characteristics (URDCs) are obtained experimentally to relate bit error probabilities to the distortion of corrupted video. Bit error rates are found first by using Viterbi's upper bounds on the bit error probability and second, by simulating nodes transmitting data spread by Total Square Correlation (TSC) codes over a Rayleigh-faded DS-CDMA channel and receiving that data using Auxiliary Vector (AV) filtering.

1. INTRODUCTION

Current wireless networking solutions do not always provide adequate support for multimedia applications because most are designed around the layered protocol architecture that forms the foundation of networking design. With the rapid growth of wireless networks, the question arises as to whether this architecture is still optimal for wireless networks, as well.[1] Wireless networks are essentially different from their wired counterparts in a number of ways. The main difference between wireless channels and wired channels is the time-varying nature of wireless channels that leads to errors due to multipath fading and cochannel interference.

In this paper, we consider a Direct Sequence-Code Division Multiple Access (DS-CDMA) visual sensor network which we assume is comprised of typically low-weight distributed sensor nodes equipped with a video camera to survey a large area. The centralized control unit at the network layer performs channel and source decoding to obtain the received video from each node. The control unit transmits information to the nodes in order to request changes in transmission parameters, such as source coding rate, channel coding rate, and transmission power. Applications of visual sensor networks include surveillance, automatic tracking and signaling of intruders within a physical area, command and control of unmanned vehicles, and environmental monitoring. Some visual sensor networks focus on image transmission and consider the trade-off between image quality and energy consumption of different routing paths.[2] Visual sensor networks that transmit video are much more challenging than typical sensor networks due to the high bit rates and delay constraints. One method is to propose a new wireless sensor node protocol stack that addresses the demanding nature of visual sensor networks.[3] But, they do not consider the improvements that can be gained with cross-layer interactions.

In our set-up, some nodes will be imaging a relatively stationary field while other nodes will be imaging scenes with a high level of motion to create a more realistic scenario where scenes with varying levels of motion exist. Video sequences

with less motion can be source encoded at a lower bit rate while still yielding good picture quality. The centralized control unit at the network layer should be able to request that the video at specific nodes be transmitted at a lower bit rate, if it is deemed as being capable of still producing adequate video quality. Nodes that compress their video at a lower bit rate can afford a higher channel coding rate and lower transmission power level so that they will cause less interference to the other nodes. For this reason, DS-CDMA is an appropriate choice for use in our visual sensor network set-up.

For video transmission over wireless channels, it is imperative to maintain a low bit error rate in order to guarantee an adequate level of viewing quality. The increased number of errors caused by transmitting over wireless channels may lead to a devastating degradation on the quality of the received video. The wireless channel's performance can be exploited by dynamically adjusting the transmission parameters through cross-layer interactions. Although Shannon's principle of separability states that it is possible to design source and channel coding separately without loss of optimality, the principle assumes that the source and channel codes are of arbitrarily long lengths. Since this assumption does not hold in practical situations due to limitations on computational power and processing delays, it is useful to consider source and channel coding jointly.[4] Hence, it is advantageous for the lower layers of the stack to consider the end application for resource management and protection strategies, and conversely, the video compression algorithms used at the application layer should consider the error protection, scheduling, etc. used at lower layers to maximize the end-to-end video quality.[5] Cross-layer design breaks the mold of the traditional layered architecture by allowing non-adjacent layers to interact with each other.

Various cross-layer schemes that have been proposed over the years. There is a joint source-channel coding method that is developed for motion-compensated Discrete Cosine Transform (DCT)-based scalable video coding and transmission.[6] Another method jointly considers source coding, channel resource allocation, and error concealment to provide a framework that balances resource consumption with end-to-end video quality in packet-based video transmission.[7] But, these methods do not consider the physical layer. The physical layer is especially important in wireless networks because it ultimately affects the performance of all other layers. It is the one layer where data is physically moved across the communications network.

Power control can be used as an indirect way of controlling a user's associated error probability in order to achieve a certain Quality-of-Service (QoS).[8] However, using power control alone is not enough to ensure adequate quality if video is transmitted. A joint source coding-power control approach exists that allocates a source coding rate and the energy-per-bit to the multiple-access interference density to each user to maximize the per-cell capacity and the end-to-end QoS for individual users.[9] There is a joint source coding and power management scheme for wireless video multicast where feedback is assumed, and Automatic Repeat Request methods are used instead of Forward Error Correction.[10] Feedback is not feasible in all communication scenarios. These techniques do not address the need for an optimal channel coding rate for each user when feedback is not possible. However, for transmission over error-prone channels, it is imperative that an acceptable choice for the channel coding rate is made to offer an adequate level of protection for the data transmitted.

There are some cross-layer designs that consider from the physical layer up through the application layer. Some jointly consider the application, data link, and physical layers in video communications, but do not consider power management.[11] [12] [13] However, a user's transmission power determines the battery life in mobile devices, the interference experienced by other users, and the network capacity. Another technique proposes a gradient-based, distortion-aware scheduling scheme for packet-based video transmission.[14] There is a cross-layer scheme that allocates the power level, source coding rate, and channel coding rate while maintaining a basic QoS to less capable receivers and an enhanced QoS for more capable receivers. They use Orthogonal Frequency-Division Multiplexing (OFDM) which is very sensitive to synchronization errors both in timing and frequency.[15]

In this work, our multi-node cross-layer optimization technique accounts for network performances all the way from the physical layer up to the application layer. At the application layer, the source coding rate for H.264/AVC video compression is determined. At the data link layer, the Rate-Compatible Puctured Convolutional (RCPC) channel coding rate is selected. And at the physical layer, the transmission power is determined. Our algorithm simultaneously allocates a source coding rate, a channel coding rate, and a power level to all nodes in a DS-CDMA visual sensor network. We recognize that different nodes of a sensor network can have different source coding rate requirements due to different scene activity and propose to jointly optimize all nodes using one of two criteria. Our first criterion results in the Minimal Average end-to-end Distortion (MAD) over all the nodes in the network while our second criterion Minimizes the Maximum Distortion (MMD) amongst all nodes. Since it would be prohibitively complex to experimentally obtain the expected distortion for each node for all possible combinations of source coding rates, channel coding rates, and power levels, we instead have chosen to

relax the optimality of the algorithm and utilize Universal Rate-Distortion Characteristics (URDCs). These characteristics show the expected distortion as a function of the bit error probabilities, P_b, after channel coding.

2. RESOURCE ALLOCATION USING CROSS-LAYER OPTIMIZATION

This work considers a wireless visual sensor network that utilizes DS-CDMA. After source and channel coding, the data is spread using a spreading code and is then carrier-modulated. But even if the spreading codes used are orthogonal to each other, transmissions from one node cause interference to the other nodes due to possible asynchronous transmissions and multipath fading. Since all nodes transmit on the same frequency, interference within a channel plays a significant role in determining the system's capacity and its quality-of-service (QoS). The transmit power for each node must be minimized to limit the interference experienced by other nodes in the system. This is critical for sensor networks, since the nodes are typically battery-operated and have limited energy. But, at the same time, a node's power must be high enough to maintain its own quality. Nodes that compress their video at a lower bit rate are left with more bits for channel coding and can afford to transmit at a lower power, thereby causing less interference to the other nodes.[16]

Assuming there are K nodes in a synchronous single-path Binary Phase Shift Keying (BPSK) channel, the received signal can be expressed as

$$\mathbf{r}(i) = A_1 b_1(i)\,\mathbf{s}_1 + \sum_{k=2}^{K} A_k b_k(i)\,\mathbf{s}_k + \mathbf{n}_k \tag{1}$$

where $A_k, b_k(i), \mathbf{s}_k, \mathbf{n}_k$, are the amplitude, symbol stream, signature sequence, and noise of node k, respectively. Ignoring thermal noise and background noise due to spurious interference allows us to assume that N_0, the noise spectral density, is entirely due to interference from other nodes in the systems

$$\sum_{k=2}^{K} A_k b_k(i)\,\mathbf{s}_k + \mathbf{n}_k \approx \sum_{k=2}^{K} A_k b_k(i)\,\mathbf{s}_k \tag{2}$$

It is a reasonable assumption that the probability distribution of the interfering nodes is a zero-mean Gaussian random variable.[17] Since user *i* has an associated power level in Watts, $S_i = E_i R_i$, the energy-per-bit to MAI ratio becomes

$$\frac{E_i}{N_0} = \frac{\frac{S_i}{R_i}}{\sum_{j\neq i}^{N} \frac{S_j}{W_t}}; i = 1, 2, 3, ..., N \tag{3}$$

where E_i is the energy-per-bit, $N_0/2$ is the two-sided noise power spectral density due to MAI in Watts/Hertz, S_i is the power of the node-of-interest in Watts, R_i is the transmitted bit rate in bits per second, S_j is the power of the interfering node in Watts, and W_t is the total bandwidth in Hertz.[9] R_i is taken to be the total bit rate used for source and channel coding. Assuming N users, R_i can be expressed as

$$R_i = \frac{R_{s,i}}{R_{c,i}}; i = 1, 2, 3, ..., N \tag{4}$$

where $R_{s,i}$ is the source coding rate for node i and $R_{c,i}$ is the channel coding rate for node i. Since $R_{s,i}$ has units of bits per second and $R_{c,i}$ is a dimensionless number, R_i will be measured in bits per second.[6]

In the first part of this work, random spreading codes are assumed and no specific receiver design is chosen. In the second part of this work, spreading codes are chosen from a set of minimum Total Squared Correlation (TSC) optimal Karystinos-Pados spreading codes[18][19] that are available for all lengths L and number of signatures K except $K = L = 4n + 1$, $n = 1, 2, \ldots$. A flow-chart showing simple algorithms based on Hadamard matrix transformations for the design of optimum binary signature sets that achieve the TSC bound for both underloaded and overloaded systems is given in.[18] After carrier demodulation, chip-matched filtering, and chip-rate sampling, auxiliary vector (AV) filtering provides MAI-suppressing despreading. Under small sample support adaptation, AV filter short-data-record estimators exhibit superior bit error rate (BER) performance in comparison to least mean squares (LMS), recursive least squares (RLS), sample matrix inversion (SMI), diagonally-loaded SMI, or multistage nested Wiener filter implementations.[20] [21]

In our visual sensor network, we assume the nodes are equipped with video cameras that monitor various fields. We assume each node has the computational power necessary for video compression. The video captured by the cameras is

compressed using the H.264/AVC video coding standard. Typically when a number of video cameras are deployed to survey a large area, some of the nodes will be imaging stationary fields whereas other nodes will be imaging scenes with more motion. Video sequences with less motion can be source encoded at a lower bit rate while still yielding good picture quality. Furthermore, it is possible for the centralized control unit to request that the video of specific source nodes be transmitted at a lower picture quality and bit rate, if it is deemed less important. We require each node to transmit at the same chip rate. So, if a node requires a lower source coding rate, a larger percentage of the total bit rate is available for channel coding. Using stronger channel coding can correct more bit errors in a link layer packet. Thus, that particular node can afford a higher bit error rate, which means that it can use a lower transmission power. This has the dual benefit that it both conserves energy for the node and reduces the interference caused to the other nodes.

In this work, we use Rate Compatible Punctured Convolutional (RCPC) codes for channel coding. Punctured convolutional codes were originally developed to simplify Viterbi decoding for rate k/n with two branches arriving at each node instead of 2^k branches. Puncturing is the process of deleting bits from the output sequence in a predefined manner so that fewer bits are transmitted than in the original code. The idea of puncturing was extended to include the concept of rate compatibility. Rate compatibility requires that a higher-rate code be a subset of a lower-rate code, or that a lower-protection code be embedded into a higher-protection code. This is accomplished by puncturing a "mother" code of rate $1/n$ to achieve higher rates. One major benefit of these RCPC codes with the same mother code is that they all can be decoded by the same Viterbi decoder.[22]

Using RCPC codes allows us to utilize Viterbi's upper bounds on the bit error probability, P_b, given by

$$P_b \leq \frac{1}{P} \sum_{d=d_{free}}^{\infty} c_d P_d \tag{5}$$

where *P* is the period of the code, d_{free} is the free distance of the code, c_d is the information error weight, and P_d is the probability that the wrong path at distance *d* is selected.[22] An AWGN channel with binary phase-shift keying (BPSK) modulation has a P_d given by

$$P_d = Q\left(\sqrt{\frac{2dR_cE_b}{N_0}}\right) \tag{6}$$

where R_c is the channel coding rate and E_b/N_0 is the energy-per-bit normalized to the single-sided noise spectral density measured in Watts/Hertz.

A centralized control unit at the network layer determines how network resources should be allocated amongst the nodes. It can request changes in transmission parameters, such as the source coding rates, channel coding rates, and transmission power levels. There are two criteria we will utilize to optimally allocate the network resources to each node in the network. The constraint for both criteria is that the chip rate be the same for all nodes. The first criterion we will employ can be formally stated as follows: Given an overall chip rate, R_{budget}, optimally allocate a source coding rate, R_s, a channel coding rate, R_c, and a power level, *S*, to all nodes such that the overall end-to-end distortion D_{ave} over all nodes is minimized

$$\min D_{ave} \text{ subject to } R_{chip} = R_{budget} \tag{7}$$

where R_{chip} is the chip rate for each node and D_{ave} is the resulting expected distortion averaged over all nodes in the system which is due to both source coding errors and channel errors. Assuming *N* nodes, D_{ave} is expressed by

$$D_{ave} = \frac{1}{N} \sum_{i=1}^{N} E\left[D_{s+c,i}\right] \tag{8}$$

where $E\left[D_{s+c,i}\right]$ is the expected distortion for user *i*. The distortion due to source coding is a result of the quantization process and is deterministic. However, the distortion due to channel errors is stochastic. Thus, the total distortion for each user is also stochastic, and we use its expected value.

The second criterion we will use to allocate resources to the nodes in the network minimizes the maximum distortion

$$\min\{\max D_{s+c,i}\} \text{ subject to } R_{chip} = R_{budget} \tag{9}$$

where R_{chip} is the chip rate for each node. Our constraint is again that the chip rate be the same for all DS-CDMA nodes. This formulation assumes that the videos from all sensors are equally important, but allows sensors, which image low-motion scenes to use a lower source coding rate. This criterion guarantees fairness among all sensors, since we are minimizing the worst distortion among all sensors. The problem is a discrete optimization problem, that is, $R_{s,i}, R_{c,i}$, and S_i can only take values from discrete sets $\mathbf{R_s}$, $\mathbf{R_c}$, and $\mathbf{S}$, respectively, i.e., $R_{s,i} \in \mathbf{R_s}$, $R_{c,i} \in \mathbf{R_c}$, $S_i \in \mathbf{S}$.[16] [6]

Since it would be prohibitively complex to experimentally obtain the expected distortion for each node for all possible combinations of source coding rates, channel coding rates, and power levels, we instead have chosen to relax the optimality of the algorithm and utilize Universal Rate-Distortion Characteristics (URDCs). These characteristics show the expected distortion as a function of the bit error probabilities, P_b, after channel coding. In the first part of this work, P_b is calculated using equations (3)-(6) for the set of source coding rates, channel coding rates and power levels. In equation (4), R_i is found for the sets of $\mathbf{R}_s$ and $\mathbf{R}_c$. These values are substituted into equation (3) for the set of $\mathbf{S}$. The results for E_i/N_0 are plugged in for E_b/N_0 in equation (6) to obtain P_d. Finally, the resulting values for P_d are substituted into equation (5) to obtain the upper bound on the P_b. This upper bound on the P_b acts as a reference for the performance of channel coding over the specified channel with the given parameters. In the second part of this work, P_b is found through simulating video transmission in a Rayleigh fading environment with interferers transmitting video data.

We calculate a RTP packet loss rate (PLR) from a certain BER, drop packets from the H.264 bitstream according to the RTP PLR, and pass the corrupted H.264 bitstream to the H.264 decoder to calculate the distortion of the uncompressed video. We assume that we know when a packet has an error. And, we manually drop packets with any errors from the H.264 encoded video stream, in accordance with the PLR_{RTP} calculated from the BER. We then calculate the distortion of this "corrupted" video stream. This creates the relation between each BER used in the URDCs and the distortion of a packet-based video stream with packet errors.

We assume the following model for the URDC for each user *i*

$$D_{s+c,i} = a\left[\log_{10}\left(\frac{1}{P_b}\right)\right]^b \tag{10}$$

where *a* and *b* are such that the square of the approximation error is minimized.[23] [24] Thus, instead of calculating the URDCs based on experimental results for every possible P_b, we instead experimentally calculate the expected distortion for a few packet loss rates associated with specific bit error rates, P_b's. We then use the model, given in equation (10), to approximate the distortion for other bit error rates. The distortion for a particular user *i*, $D_{s+c,i}$, given a particular source coding rate, $R_{s,i}$, is a function of the bit error rate. Therefore, URDCs will give a family of $D_{s+c,i}$ versus $1/P_b$ curves given a set of source coding rates for each type of node.[23] [24]

We perform the optimization procedure using the proposed model for URDCs. The data points used to obtain the parameters *a* and *b* are obtained by corrupting the video stream with packet errors based on a calculated P_b, decoding the corrupted video bit stream with the H.264/AVC codec, calculating the distortion, repeating this experiment 300 times and then taking the average distortion. We assume that there are two possible motion levels viewed by the sensor nodes, low motion and high motion. The "Akiyo" sequence is used to represent a low-motion node, and the "Foreman" sequence is used to represent a high-motion node. It is necessary to have two sets of URDC curves, one for each level of motion. The characteristics were obtained for both video sequences at a frame rate of 15 f/s.

We use BPSK modulation and RCPC codes with mother code rate 1/4 for channel coding.[22] We first assume random spreading codes with the same processing gain for all nodes, so our constraint that the chip rate be the same for all DS-CDMA nodes translates into a constraint on the transmitted bit rate given in (4). We examine various target chip rate constraints at 192000 c/s, 144000 c/s, 96000 c/s, and 48000 c/s. The set of admissible source coding rates and corresponding channel coding rates for the various chip rates are

$$R_{chip} = 192000c/s \rightarrow \mathbf{R}_s, \mathbf{R}_c \in \{(64kbps, 1/3), (96kbps, 1/2), (128kbps, 2/3)\} \tag{11}$$

$$R_{chip} = 144000c/s \rightarrow \mathbf{R}_s, \mathbf{R}_c \in \{(48kbps, 1/3), (64kbps, 1/2), (96kbps, 2/3)\} \tag{12}$$

$$R_{chip} = 96000c/s \rightarrow \mathbf{R}_s, \mathbf{R}_c \in \{(32kbps, 1/3), (48kbps, 1/2), (64kbps, 2/3)\} \tag{13}$$

$$R_{chip} = 48000c/s \rightarrow \mathbf{R}_s, \mathbf{R}_c \in \{(16kbps, 1/3), (24kbps, 1/2), (32kbps, 2/3)\} \tag{14}$$

The power levels in Watts were chosen from $\mathbf{S} \in \{5, 10, 15\}$. The total bandwidth, W_t, was set to 20MHz.

In Tables 1-4, we show how the network resources should be assigned for various distributions of the two types of nodes for different target chip rates. The low-motion nodes' source coding rate in bits per second, channel coding rate, and power level in Watts are represented by R_{s1}, R_{c1}, and S_1, respectively, and the high-motion nodes' parameters are represented by R_{s2}, R_{c2}, and S_2.The number of low-motion nodes is given under column, "Low", and the number of high-motion nodes is given under column,"High". "MAD" corresponds to the the method of Minimizing the Average end-to-end Distortion over all users, and "MMD" corresponds to the technique of Minimizing the Maximum Distortion. In Tables 1-4, the distribution of the two types of nodes is varied while the total number of nodes is kept constant. Tables 1 and 2 use the MAD criterion while Tables 3 and 4 utilize the MMD criterion. There are equal numbers of the nodes viewing low-motion scenes and nodes viewing high-motion scenes. We give the resulting average end-to-end peak signal-to-noise ratio, PSNR, in dB for the entire network as a measure of performance for the MAD experiments. We also use the minimum PSNR as a measure of performance for the MMD experiments. The PSNR is calculated from the expected distortion

$$PSNR = 10 \log \left(\frac{255 * 255}{E\{D_{s+c}\}} \right) \tag{15}$$

where PSNR is the peak signal-to-noise ratio and $E\{D_{s+c}\}$ is the expected distortion due to source and channel coding

Table 1. MAD with Equal Distributions of Node Types: Target chip rate = 144000 c/s

Low	(R_{s1}, R_{c1}, S_1)	$D_{s+c,1}$	High	(R_{s2}, R_{c2}, S_2)	$D_{s+c,2}$	D_{ave}	$PSNR_{ave}$
10	$(96k, 2/3, 15)$	1.8	10	$(96k, 2/3, 15)$	12.1	6.9	$39.7dB$
30	$(96k, 2/3, 10)$	4.7	30	$(96k, 2/3, 15)$	16.0	10.4	$38.0dB$
50	$(48k, 1/3, 5)$	11.9	50	$(96k, 2/3, 10)$	20.6	16.3	$36.0dB$
70	$(48k, 1/3, 5)$	20.0	70	$(96k, 2/3, 10)$	32.8	26.4	$33.9dB$
90	$(48k, 1/3, 10)$	24.5	90	$(48k, 1/3, 15)$	68.3	46.4	$31.5dB$

Table 2. MAD with Equal Distributions of Node Types: Target chip rate = 96000 c/s

Low	(R_{s1}, R_{c1}, S_1)	$D_{s+c,1}$	High	(R_{s2}, R_{c2}, S_2)	$D_{s+c,2}$	D_{ave}	$PSNR_{ave}$
10	$(64k, 2/3, 15)$	3.0	10	$(64k, 2/3, 15)$	22.4	12.7	$37.1dB$
30	$(48k, 1/2, 5)$	7.9	30	$(64k, 2/3, 15)$	23.5	15.7	$36.2dB$
50	$(48k, 1/2, 5)$	9.7	50	$(64k, 2/3, 10)$	35.1	22.4	$34.6dB$
70	$(48k, 1/2, 10)$	11.7	70	$(48k, 1/2, 15)$	49.7	30.7	$33.3dB$
90	$(32k, 1/3, 10)$	19.8	90	$(48k, 1/2, 15)$	58.9	39.3	$32.2dB$

Table 3. MMD with Equal Distributions of Node Types: Target chip rate = 144000 c/s

Low	(R_{s1}, R_{c1}, S_1)	$D_{s+c,1}$	High	(R_{s2}, R_{c2}, S_2)	$D_{s+c,2}$	D_{ave}	$PSNR_{ave}$
10	$(72k, 1/2, 15)$	1.8	10	$(96k, 2/3, 15)$	12.1	6.9	$39.7dB$
30	$(48k, 1/3, 5)$	9.6	30	$(96k, 2/3, 15)$	14.5	12.1	$37.3dB$
50	$(48k, 1/3, 5)$	17.9	50	$(96k, 2/3, 15)$	18.9	18.4	$35.5dB$
70	$(48k, 1/3, 5)$	20.0	70	$(96k, 2/3, 10)$	32.8	26.4	$33.9dB$
90	$(48k, 1/3, 10)$	24.5	90	$(48k, 1/3, 15)$	68.3	46.4	$31.5dB$

We see that in most MAD cases, high-motion nodes are assigned a higher source coding rate than the low-motion nodes. This is because the drop in the end-to-end distortion when increasing the source coding rate for a high-motion video sequence is more significant than the effect of employing stronger channel coding. However, the distortions for the low-motion video sequence remain relatively low even when the source coding rate is decreased, so it can afford to transmit at a lower source coding rate in some cases. Since low-motion video sequences are more robust to errors, low-motion nodes are also assigned a lower power than the high-motion nodes in most cases even when they both have the same source and channel coding rates. The times when all of the source coding rates, channel coding rates, and power levels are the same is usually when the network resources are being strained with a large amount of nodes transmitting at high rates with limited

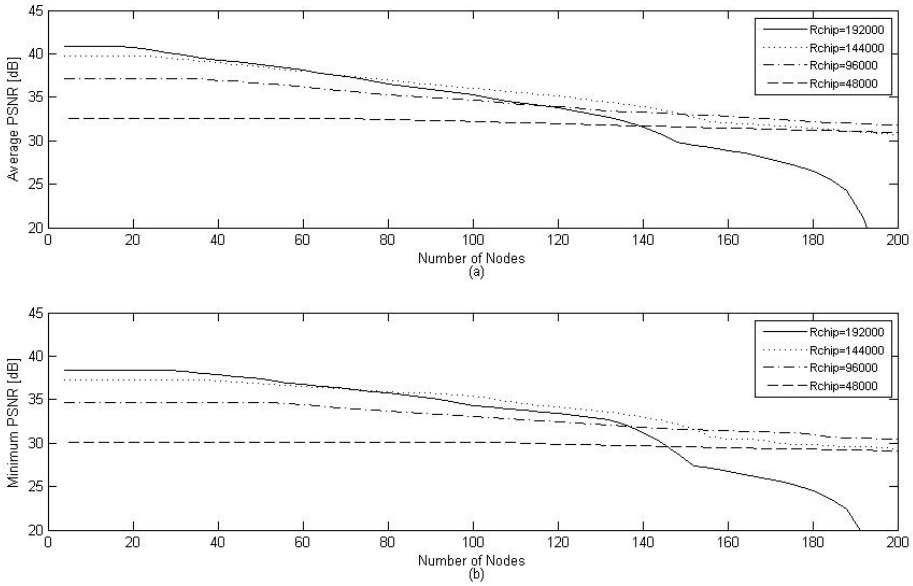
Table 4. MMD with Equal Distributions of Node Types: Target chip rate = 96000 c/s

Low	(R_{s1}, R_{c1}, S_1)	$D_{s+c,1}$	High	(R_{s2}, R_{c2}, S_2)	$D_{s+c,2}$	D_{ave}	$PSNR_{ave}$
10	$(48k, 1/2, 15)$	3.0	10	$(48k, 1/2, 15)$	22.4	12.7	$37.1dB$
30	$(48k, 1/2, 5)$	7.9	30	$(64k, 2/3, 15)$	23.5	15.7	$36.2dB$
50	$(48k, 1/2, 5)$	14.6	50	$(64k, 2/3, 15)$	32.2	23.4	$34.4dB$
70	$(32k, 1/3, 5)$	25.7	70	$(64k, 2/3, 15)$	42.2	34.0	$32.8dB$
90	$(32k, 1/3, 5)$	51.5	90	$(48k, 1/2, 15)$	50.5	51.0	$31.1dB$

bandwidth. It is forced to assign the lower source coding rates and power levels to both types of nodes in those cases. We note that overloading the network results in a drastic drop in the system performance to the point of obtaining unrealistic distortions that would result in no viewable video. We see that the system starts to deteriorate when there are more than 150 total nodes.

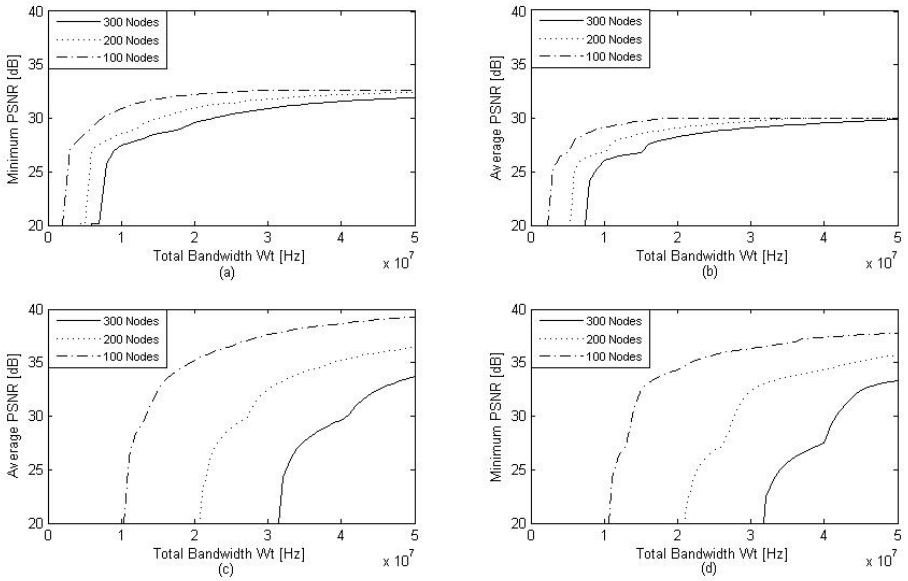
In Figure 1, we show graphs of the resulting average PSNR or the minimum PSNR of the overall network for various target chip rates using the two criteria. We plot the average PSNR against the number of nodes in the network for all the target chip rates in Figure 1 (a). There are equal numbers of the low-motion nodes and the high-motion nodes. In Figure 1 (b), we plot the minimum PSNR versus the number of nodes in the network for different target chip rates. The minimum PSNR corresponds to the maximum distortion which is minimized according to the MMD criteria. Even though initially the higher target chip rates result in the higher PSNR, we see that in both figures the downward slope of the curve for higher target chip rates is also steeper. This translates into the network being overloaded sooner with higher chip rates.

Figure 1. (a) Average PSNRs versus Number of Nodes for different Target Chip Rates [c/s] using MAD (b) Minimum PSNRs versus Number of Nodes for different Target Chip Rates [c/s] using MMD



In Figures 2 (a) and (c), we plot the average PSNR using MAD versus the total bandwidth, W_t, for the target chip rates of 48000 c/s and 192000 c/s respectively. We plot the minimum PSNR using MMD versus W_t in Figures 2 (b) and (d). Since these figures show curves for different numbers of total nodes, they can be used to determine what total bandwidth is required to accommodate a desired number of nodes and a desired average end-to-end video quality level. We see the drastic drop in performance at low W_t. We also observe the relationship between the number of nodes and the required W_t to handle that number of nodes for a certain target chip rate.

Figure 2. (a) Average PSNRs using MAD versus Total Bandwidth for Target Chip Rate = 48000 c/s (b) Minimum PSNRs using MMD versus Total Bandwidth for Target Chip Rate = 48000 c/s (c) Average PSNRs using MAD versus Total Bandwidth for Target Chip Rate = 192000 c/s (d) Minimum PSNRs using MMD versus Total Bandwidth for Target Chip Rate = 192000 c/s



For comparison sake, Figure 3 show similar graphs with various target chip rates for the 50-node case. As expected at the higher target chip rates, it takes more bandwidth to reach the maximum level of PSNR. We see how the improvement in average PSNR between target chip rates 48000 c/s and 96000 c/s is significantly greater than the improvement between target chip rates 144000 c/s and 192000 c/s, even though there is the same 48000 c/s difference in chip rate. If there are bandwidth restrictions on the network, one can use these figures to determine the target chip rate that will result in the best average end-to-end PSNR overall for a given total bandwidth.

The results up until this point assumed random spreading codes, that interference was seen as white Gaussian noise, and no specific receiver design. Now we shall consider more specific parameters such as Total Squared Correlation (TSC) codes, interleaving, a Rayleigh fading channel, and Auxiliary Vector (AV) filtering at the receiver for the proposed optimization method to validate the theoretical assumptions made in the first part of this work. Instead of utilizing the Viterbi upper bound on the probability of error, we find the probabilities of error through simulating actual simultaneous node transmissions. We transmit multiple nodes' data, divided evenly between high-motion nodes and low-motion nodes. We spread the data with TSC codes, encode the spread data with RCPC codes, and interleave the encoded data. Next, we send the data over a Rayleigh fading channel with 3 multipaths. At the receiver, the data is demodulated and despread using the AV filter. The despread data is channel decoded with a Viterbi decoder and then passed through a deinterleaver. After repeated runs, the probability of error is determined. In Figure 4, the *RSC-AWGN* method refers to using random spreading codes, interference as white Gaussian noise, no specific receiver and hence, utilizing Viterbi upper bound on the probability of error. The *TSC-AV* method refers to simulating transmission while incorporating TSC spreading codes, interleaving, a Rayleigh fading channel, and AV filtering at the receiver. Utilizing the *TSC-AV* method results in higher average PSNR values for very low and very high numbers of nodes for MAD, and hence better average end-to-end video quality even when using a more realistic Rayleigh fading channel. The MMD curves follow each other's behavior closely. *TSC-AV* uses spreading codes with an interference suppression algorithm. Since the *RSC-AWGN* method uses random codes that have no interference suppression characteristics, all of the power transmitted by other nodes arrives as interference at the node-of-interest, so we see its performance is a bit less at some points.

Figure 3. (a) Average PSNRs using MAD versus Total Bandwidth for 50 Nodes for different Target Chip Rates [c/s] (b) Minimum PSNRs using MMD versus Total Bandwidth for 50 Nodes for different Target Chip Rates [c/s]

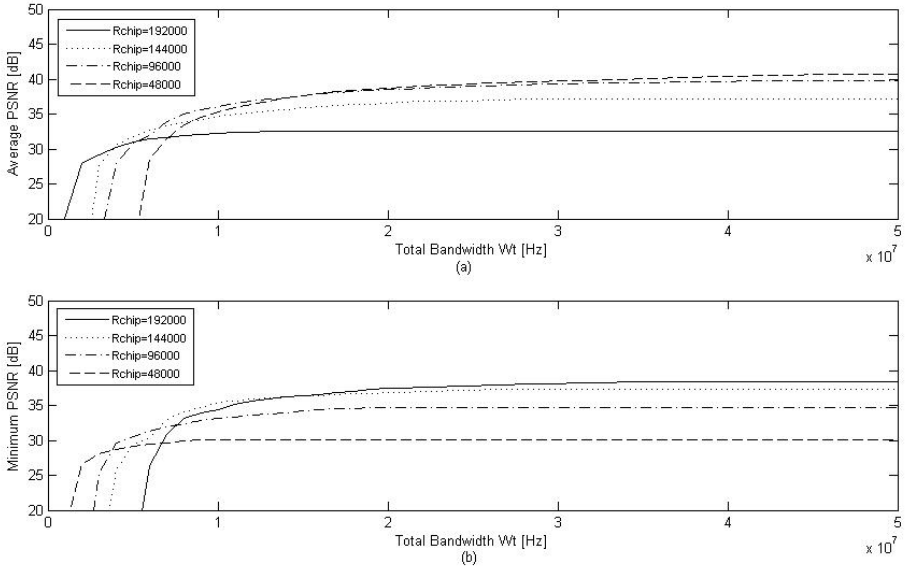
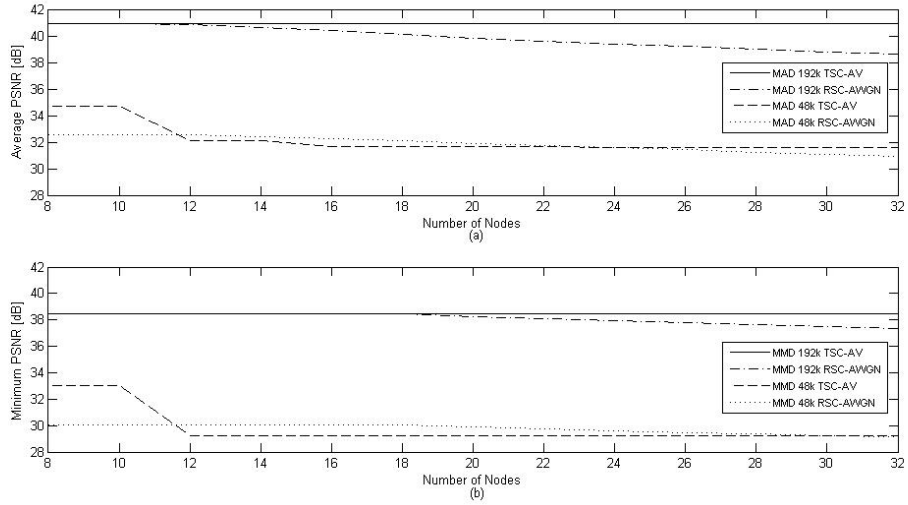


Figure 4. (a) Comparing RSC-AWGN and TSC-AV methods using MAD for different Target Chip Rates [c/s] (b) Comparing RSC-AWGN and TSC-AV methods using MMD for different Target Chip Rates [c/s]



3. CONCLUSION

In this paper, we present a cross-layer optimization algorithm that works across the physical layer, the data link layer, and the application layer in a wireless visual sensor network. This algorithm accounts for network performances all the way from the physical layer up to the application layer. At the application layer, we determined the source coding rate, R_s, for video compression. At the data link layer, we assign the channel coding rate, R_c. At the physical layer, we select the transmission power level, S. The algorithm shows how to distribute these parameters among all the nodes transmitting in the network. To create a realistic DS-CDMA visual sensor network, different levels of motion are assumed to be imaged by the nodes. By utilizing the parametric model for the URDCs, we find each node's expected distortion for only the probabilities of error calculated for a small number of source coding rates, channel coding rates, and power levels and used the model to estimate the distortion for other probabilities of error. This reduced the computational complexity of

the solution significantly. We present the combinations of $\{R_s, R_c, S\}$ for each node that result in the minimal average end-to-end distortion over all nodes in the system and the combinations that minimize the maximum distortion. We also show how to determine the minimum total bandwidth needed to obtain a specific level of quality for the desired number of nodes and which target chip rate achieves the highest average PSNR for a given total bandwidth.

REFERENCES

[1] V. T. Raisinghani and S. Iyer, "Cross-layer feedback architecture for mobile device protocol stacks," *IEEE Communications Magazine*, vol. 44, no. 1, pp. 85–92, January 2006.

[2] K.-Y. Chow, K.-S. Lui, and E. Y. Lam, "Balancing Image Quality and Energy Consumption in Visual Sensor Networks," in *2006 1st International Symposium on Wireless Pervasive Computing*, January 2006.

[3] T. Zahariadis and S. Voliotis, "Open issues in Wireless Visual Sensor Networking," in *14th International Workshop on Systems, Signals, and Image Processing, 2007 and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services*, June 2007, pp. 335–338.

[4] K. Ramchandran and M. Vetterli, *Multiresolution Joint Source-Channel coding for Wireless Channels,* chapter in Wireless Communications: A Signal Processing Perspective, V. Poor and G. Wornell ed. Prentice-Hall, 1998.

[5] M. van der Schaar and D. S. Turaga, "Cross-Layer Packetization and Retransmission Strategies for Delay-Sensitive Wireless Multimedia Transmission," *IEEE Transactions on Multimedia*, vol. 9, no. 1, pp. 185–197, January 2007.

[6] L. P. Kondi, F. Ishtiaq, and A. K. Katsaggelos, "Joint Source-Channel Coding for Motion-Compensated DCT-Based SNR Scalable Video," *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 1043–1052, September 2002.

[7] A. K. Katsaggelos, Y. Eisenberg, F. Zhai, R. Berry, and T. N. Pappas, "Advances in efficient resource allocation for packet-based real-time video transmission," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 135–147, January 2005.

[8] L. C. Yun and D. G. Messerschmitt, "Variable quality of service in CDMA systems by statistical power control," in *Proc. ICC*, vol. 2, Seattle, WA, 1995, pp. 713–719.

[9] Y. S. Chan and J. W. Modestino, "A Joint Source Coding-Power Control Approach for Video Transmission Over CDMA Networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1516–1525, December 2003.

[10] I. V. Bajic, "Efficient Cross-Layer Error Control for Wireless Video Multicast," *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 276–285, March 2007.

[11] Y. Andreopoulos, N. Mastronarde, and M. V. D. Schaar, "Cross-Layer Optimized Video Streaming Over Wireless Multihop Mesh Networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 11, pp. 2104–2115, November 2006.

[12] L. P. Kondi, D. Srinivasan, D. A. Pados, and S. N. Batalama, "Layered Video Transmission over Multirate DS-CDMA Wireless Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1629–1637, December 2005.

[13] Y. Shen, P. C. Cosman, and L. B. Milstein, "Error-resilient Video Communications Over CDMA networks With A Bandwidth Constraint," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3241–3252, November 2006.

[14] P. Pahalawatta, R. Berry, T. Pappas, and A. Katsaggelos, "Content-Aware Resource Allocation and Packet Scheduling for Video Transmission over Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 4, pp. 749–759, May 2007.

[15] Y. S. Chan, J. W. Modestino, Q. Qu, and X. Fan, "An End-to-End Embedded Approach for Multicast/Broadcast of Scalable Video over Multiuser CDMA Wireless Networks," *IEEE Transactions on Multimedia*, vol. 9, no. 3, pp. 655–667, April 2007.

[16] E. S. Pynadath and L. P. Kondi, "Cross-Layer Optimization with Power Control in DS-CDMA Visual Sensor Networks," in *Proc. of the IEEE International Conference on Image Processing*, Atlanta, GA, 2006, pp. 25–28.

[17] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, J. L. A. Weaver, and C. E. W. III, "On the Capacity of a Cellular CDMA System," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 2, pp. 303–312, May 1991.

[18] G. N. Karystinos and D. A. Pados, "New Bounds on the Total Squared Correlation and Optimum Design of DS-CDMA Binary Signature Sets," in *IEEE Transactions On Communications*, vol. 51, Jan 2003, pp. 48–51.

[19] C. Ding, M. Golin, and T. Klove, "Meeting the Welch and Karystinos-Pados Bounds on DS-CDMA Binary Signature Sets," The Hong Kong University of Science and Technology," HKUST Theoretical Computer Science Research Report, 2003.

[20] A. Kansal, , S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering and DS-CDMA multipath fading channels," *IEEE Journal on Selected Areas of Communications*, vol. 16, pp. 1765–1773, December 1998.

[21] J. S. Goldstein, I. S. Reed, and L. L. Sharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Transactions on Image Processing*, vol. 44, pp. 2943–2959, November 1998.

[22] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Transactions on Communications*, vol. 36, no. 4, pp. 389–400, April 1988.

[23] L. P. Kondi and A. K. Katsaggelos, "Joint Source-Channel Coding for Scalable Video Using Models of Rate-Distortion Functions," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, Utah, 2001, pp. 1377–1380.

[24] M. Bystrom and T. Stockhammer, "Modeling of operational distortion-rate characteristics for joint source-channel coding of video," in *Proc. of the IEEE International Conference on Image Processing*, Kobe, Japan, 1999, pp. 359–362.